u

UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/815,591	03/23/2001	Victor Spivak	DOCM-002/01US	5572

			EXAMINER
23419	7590	06/03/2004	BASEHOAR, ADAM L

COOLEY GODWARD, LLP
3000 EL CAMINO REAL
5 PALO ALTO SQUARE
PALO ALTO, CA 94306

ART UNIT	PAPER NUMBER
2178	

DATE MAILED: 06/03/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

	Application No.	Applicant(s)
Office Action Summary	09/815,591	SPIVAK ET AL.
	Examiner	Art Unit	
	Adam L Basehoar	2178	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *23 March 2001*.

2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1-20* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-20* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☒ The drawing(s) filed on *23 March 2001* is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a) ☐ All b) ☐ Some * c) ☐ None of:

      1. ☐ Certified copies of the priority documents have been received.

      2. ☐ Certified copies of the priority documents have been received in Application No. _____.

      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
    Paper No(s)/Mail Date **01/28/03 01/31/02 06/18/01**

4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

DETAILED ACTION

1.   This action is responsive to communications: The Application filed on 03/23/01, which

claims priority to a provisional application, filed 03/27/00 and the three IDS filed on 06/18/01,

01/31/02, and 01/28/03 respectively.

2.   Claims 1-20 are pending in the case. Claims 1, 7, and 14 are independent claims.

Claim Rejections - 35 USC § 103

3.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in
> section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are
> such that the subject matter as a whole would have been obvious at the time the invention was made to a person
> having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the
> manner in which the invention was made.

4.   Claims 1-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Smadja (US:

6,621,930 09/16/03).

-In regards to independent claims 1, 7, and 14, Smadja teaches a computer-implemented

method comprising a processor (Fig. 3: 32) and memory connected to said processor

(Fig. 3: 44), wherein the method further comprises;

recognizing a concept (frequency statistics of token words)(column 3, lines 27-

42) that represents a basic idea ("glove" "bat" "single") in a document format (column 4,

lines 61-65; and

incorporating said concept in a concept model (baseball category)(column 4, lines 61-65).

Smadja further teaches wherein the document format could be any number of common document formats including an electronic email message, a word processing document, hypertext document, and any number of other types of documents (columns 3 & 4, lines 23-26 & 51-53). Smadja does not teach wherein the initial document format have to be converted to one of the common document formats to be processed. It would have been obvious to one of ordinary skill in the art at the time of the invention for Smadja to have converted initial format document to one of the common document formats listed above, because if the initial document was not in a format recognized by the categorization facility (Fig. 3: 48), the initial document would not be able to be categorized into one of the many document categories.

-In regard to dependent claims 2 and 8, Smadja teaches identifying a plurality of features (tokens in a lexicon list: "glove" "bat" "single")(column 4, lines 61-62) in said document format, wherein said plurality of features represent evidence ("helps distinguish one category from another")(column 4, lines 45-54) of said concept in said format.

-In regard to dependent claims 3 and 9, Smadja teach calculating a concept weight for said concept (frequency of occurrence of a plurality of token) using a plurality of feature weights (token frequencies in initial document) associated with said plurality of features (tokens), wherein said concept weight represents a recognition confidence level for said concept (column 3, lines 4-22); and

comparing said concept weight with a predetermined thresholds (columns 3 & 5, lines 36-42 & 23-27).

-In regard to dependent claims 4, 11, 13, and 19, Smadja teaches by referencing said concept model (statistical category), generating an auto-attribute/category (column 4, lines 11-23)(Fig. 4), said auto-attribute/category being descriptive label (i.e. baseball, neutral, SPAM, etc.)(Fig. 7: 116, 118, 80) for said format/category.

-In regard to dependent claims 5, 12, 18, and 20, Smadja teaches by referencing said concept model (statistical category), assigning said document format to a subject category/modeling directory in a categorization taxonomy (Fig. 4: 78 & 80) including a plurality of categories (Baseball, Java, C++, Neutral)(column 4, line 61 & Fig. 7: 116 & 118).

-In regard to dependent claim 6, Smadja do teach wherein a common document format was hypertext or other like documents (column 4, lines 52-53). Smadja does not specifically teach wherein a common format was an XML document. It would have been obvious to one of ordinary skill in the art at the time of the invention, for one of the common formats of Smadja to have been XML, because XML was notoriously well known to be synonymous with hypertext documents, as well as being an International document standard, and well known for its separation of data content which was the embodiment of the Smadja reference. It also would have been obvious to one of ordinary skill in the art at the time of the invention for Smadja to have converted initial format document to an XML common document format, because if the initial document was not in a format recognized by the categorization facility (Fig. 3: 48), the initial document would not be able to be categorized into one of the many document categories.

-In regard to dependent claim 10, Smadja teaches incorporating said recognition confidence level (category threshold) in said conceptual model (category) based on the training data (column 3, lines 36-42)

In regard to dependent claim 15, as shown above, Smadja teaches wherein the document format could be any number of common document formats including an electronic email message, a word processing document, hypertext document, and any number of other types of documents (columns 3 & 4, lines 23-26 & 51-53). Smadja does not teach wherein the initial document format have to be converted to one of the common document formats to be processed. It would have been obvious to one of ordinary skill in the art at the time of the invention for Smadja to have converted initial format document to one of the common document formats listed above, because if the initial document was not in a format recognized by the categorization facility (Fig. 3: 48), the initial document would not be able to be categorized into one of the many document categories.

-In regard to dependent claim 16, Smadja teaches separating the text content from said initial format document for categorizing documents based on statistical techniques (column 4, lines 1-6). As shown above in dependent claim 15, Smadja does not teach converting the initial document format into a common document format. It would have been obvious to one of ordinary skill in the art at the time of the invention for Smadja to have converted initial format document to one of the common document formats listed above, because if the initial document was not in a format recognized by the categorization facility (Fig. 3: 48), the initial document would not be able to be categorized into one of the many document categories;

wherein it would have also been obvious to incorporate the text from the initial document into the said common document, because Smadja teaches the textual content was what was needed to categorize the incoming documents (column 4, lines 1-6).

-In regard to dependent claim 17, Smadja teaches identifying a plurality of features (tokens in a lexicon list: "glove" "bat" "single")(column 4, lines 61-62) in said document format, wherein said plurality of features represent evidence ("helps distinguish one category from another")(column 4, lines 45-54) of said concept in said format. Smadja further teaches calculating a concept weight for said concept (frequency of occurrence of a plurality of token) using a plurality of feature weights (token frequencies in initial document) associated with said plurality of features (tokens), wherein said concept weight represents a recognition confidence level for said concept (column 3, lines 4-22); and

comparing said concept weight with a predetermined thresholds (columns 3 & 5, lines 36-42 & 23-27).

Conclusion

5.   The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

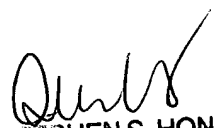|		
US: 6,101,515	08/08/00	Wical et al.
US: 6,442,545	08/27/02	Feldman et al.
US: 6,675,162	01/06/04	Russell-Falla et al.

6.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Adam L Basehoar whose telephone number is (703) 305-7212.

The examiner can normally be reached on M-F: 7:30am - 4:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Heather Herndon can be reached on (703) 308-5186. The fax phone number for the

organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent

Application Information Retrieval (PAIR) system. Status information for published applications

may be obtained from either Private PAIR or Public PAIR. Status information for unpublished

applications is available through Private PAIR only. For more information about the PAIR

system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR

system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ALB

STEPHEN S. HONG
PRIMARY EXAMINER